



## Chapter 3

### Forecasting the South African Economy:

### A DSGE-VAR Approach

#### 3.1 Introduction

The controversy about methods for evaluating the empirical relevance of economic models is not new. However, two distinct approaches has emerged since the early 1980s. First, the standard econometric approach in which an economic model should be embedded within a complete probability model and analyzed using statistical methods (Watson, 1993). For instance, Vector Autoregression (VAR) models introduced by Sims (1980), which can be taken directly to the data to perform statistical hypothesis. VAR models also became popular in the forecasting literature pioneered by Litterman (1986b). Although VAR models

have been proved to be reliable tools in terms of data description and forecasting, they are subject to Lucas critique (Lucas, 1976) and also fail to take account of nonlinearities in the economy.

The second approach, pioneered by Kydland and Prescott (1982) and Long and Plosser (1983), has become increasingly popular for evaluating dynamic macroeconomic models. Dynamic stochastic general equilibrium (DSGE) models are explicitly derived from the first principles. DSGE models describe the general equilibrium of a model economy in which agents like consumers and firms maximize their objectives subject to budget and resource constraints (Del Negro and Schorfheide, 2003). Therefore, the DSGE structural (or 'deep') parameters, in principle, do not vary with the policy regime. However, the calibrated DSGE models are typically too stylized to be taken directly to the data and often yield fragile results (Stock and Watson, 2001; Ireland, 2004).

In this paper, we develop an estimated DSGE model for forecasting the Gross National Product (GNP), consumption, investment and hours worked for South African economy. Our proposed hybrid DSGE-VAR model combines a micro-founded DSGE model with the flexibility of a VAR framework. The model is estimated using maximum likelihood technique based on quarterly data obtained from the South African Reserve Bank over the period of 1970:1-2000:4. Based on a recursive estimation using the Kalman filter algorithm, the out-of-sample forecasts from the hybrid model are then compared with the forecasts generated

from the Classical and Bayesian variants of the VAR for the period 2001:1-2005:4.

The remainder of the paper is organized as follows. Section 2 lays out the theoretical model, while Section 3 describes the hybrid model. Results are presented in Section 4 and Section 5 concludes.

### 3.2 The Model Economy

The model economy, here, is based on the benchmark real business cycle model developed by Hansen (1985). Equilibrium models have been criticized for depending heavily on individuals' substitution of leisure and work responding to the change in interest rate or wage. Hansen (1985) and Rogerson (1988) argue that in the real economy labor is indivisible. Individuals either work full time or not at all. Other features of Hansen's indivisible labor are exactly the same as standard real business cycle models, such as Kydland and Prescott (1982). The economic environment is described below.

The model economy is populated by infinitely-lived households. The preferences of households are assumed to be identical. Households maximize expected life-time utility:

$$U(C_t, H_t) = E_t \sum_{t=0}^{\infty} \beta^t (\ln C_t - \gamma H_t), \quad 0 < \beta < 1 \quad \gamma > 0 \quad (1)$$

where  $C_t$  and  $H_t$  are consumption and hours worked respectively,  $\beta$  is the discount factor that households apply to future utility.

The technology is defined as a standard Cobb-Douglas production function with constant-returns-to-scale:

$$Y_t = Z_t K_{t-1}^\rho (\eta^t H_t)^{1-\rho}, \quad 0 < \rho < 1 \quad \eta > 1 \quad (2)$$

where  $\rho$  is the fraction of household's income that goes to the capital input and  $1 - \rho$  is the fraction that goes to the labor input.  $\eta$  measures the gross rate of labor-augmenting technological process.  $Z_t$  is the technology shock, which is exogenously evolving according to the law of motion:

$$\log Z_t = (1 - \psi) \log \bar{Z} + \psi \log Z_{t-1} + \epsilon_t, \quad \epsilon_t \sim i.i.d.(0, \sigma^2) \quad (3)$$

where  $\psi$  and  $\bar{Z}$  are parameters, and  $0 < \psi < 1$ . The innovation  $\epsilon_t$  is normally distributed.

As in a neoclassical growth model, capital stock depreciates at a constant rate of  $\delta$ , and households invest a fraction of income in capital stock in each period. This amount of investment forms part of productive capital in current period. Therefore the law of motion for aggregate capital stock is

$$K_{t+1} = (1 - \delta)K_t + I_t, \quad 0 < \delta < 1 \quad (4)$$

The model economy is a closed economy, where  $Y_t = C_t + I_t$ . In equilibrium the representative consumer maximizes his or her utility function (1) subject to the aggregate constraints

$$Y_t = C_t + I_t$$

$$Y_t = Z_t K_{t-1}^\rho (\eta^t H_t)^{1-\rho}$$

$$K_{t+1} = (1 - \delta)K_t + I_t$$

$$\log Z_t = (1 - \psi)\log \bar{Z} + \psi \log Z_{t-1} + \epsilon_t, \quad \epsilon_t \sim i.i.d.(0, \sigma_\epsilon^2)$$

### 3.3 The Hybrid Model: A DSGE-VAR Approach

Kydland and Prescott (1982) argue that in the basic RBC framework, the U.S. business cycle fluctuations are purely driven by real technology shocks. This one-shock assumption makes real business cycle models stochastically singular. Using a version of the King *et al.* (1988) model, Ingram *et al.* (1994) point out that it is impossible to derive the realizations of the productivity shocks using a singular model if the variance-covariance matrix of the observable variables is actually nonsingular. In order to overcome this singularity problem, Ingram *et al.* (1994), DeJong *et al.* (2000a, b), Ireland (2001 and 2002), and Kim (2000) elaborate the DSGE model to a more elaborate model by including as many shocks as there are endogenous variables in the model. This approach, in addition, can be served to identify sources of output variation<sup>1</sup>.

Recently, Ingram and Whiteman (1994), DeJong *et al.* (2000a, b), and Schorfheide (2000) have used a Bayesian framework to estimate and evaluate

---

<sup>1</sup>The literature suggest that the technology shocks are primarily responsible for the postwar U.S. business cycle fluctuations.

DSGE models. The principle underlying a Bayesian analysis of DSGE models is to combine prior and likelihood functions in order to obtain posterior distributions of the variables of interest. However, different methods have been applied to this kind of research. Ingram and Whiteman (1994) use the King *et al.* (1988) real business cycle model as a source of priors in Bayesian VAR (BVAR) forecasting exercises, whereas, the method pursued by DeJong *et al.* (2000a, b) and Schorfheide (2000) lies between calibration and maximum likelihood estimation exclusively within the DSGE model. Moreover, there is a significant progress in the development of DSGE models that deliver acceptable forecasts (Smets and Wouters, 2003a, b, 2004; Del Negro and Schorfheide, 2004, Del Negro *et al.*, 2005). The authors use prior information derived from DSGE models in the estimation of the VARs. The hybrid models are then used to perform forecasting exercises. The empirical results suggest that the out-of-sample forecasts from the estimated DSGE models outperform the VARs estimated with simple least squares methods.

The approach proposed in this paper is based on Ireland (2004), which is different from the ones discussed above. We augment the linearized solution of the model with unobservable errors that have a VAR representation. This approach was developed originally by Sargent (1989) and pursued by Altug (1989), Watson (1993), Hall (1996), and McGrattan *et al.* (1997). The hybrid DSGE-VAR model is constructed as follows.

The approximated solution is applied to the log-linearized model<sup>2</sup>, where a serially correlated residual is augmented to each equation as in (5)

$$\hat{\pi}_t = A\hat{x}_t + \mu_t \tag{5}$$

and

$$\hat{x}_t = B\hat{x}_{t-1} + C\epsilon_t \tag{6}$$

$$\mu_t = D\mu_{t-1} + \xi_t \quad \xi_t \sim i.i.d.(0, \sigma_\xi^2) \tag{7}$$

where  $\hat{\pi}_t$  is the vector of all de-trended endogenous variables in log-deviations,  $\hat{\pi}_t = [\hat{y}_t \ \hat{c}_t \ \hat{i}_t \ \hat{h}_t]'$ , and  $\hat{x}_t$  is the vector of de-trended state variables in log-deviations,  $\hat{x}_t = [\hat{k}_t \ \hat{z}_t]'$ . The matrix  $D$  is governing the persistence of the VAR residuals. The covariance matrix of the residuals in (7),  $E\xi_t\xi_t' = V$ , is uncorrelated with the innovation to technology,  $\epsilon_t$ . The covariance matrix  $V$  is also constrained to be positive definite (Hamilton, 1994: 147).

Sargent (1989) assumes the measurement errors are uncorrelated with the data generated from the model by restricting  $D$  and  $V$  matrices as diagonal. In this paper, however, we estimate the DSGE model both with and without the restrictions on  $D$  and  $V$  matrices. The advantage of imposing no restrictions on  $D$  and  $V$  matrices is that the residuals in  $\mu_t$  can capture not only the measurement errors, but also the movements and co-movements in the data that the stylized real business cycle model cannot explain (Ireland, 2004: 1210). Furthermore, in

---

<sup>2</sup>Appendix B describes the steady state of the model as well as the the log-linearized model

order to guarantee the residuals in  $\mu_t$  are stationary, the eigenvalues of the matrix  $D$ , which govern the persistence of the VAR residuals, are constrained to be less than one.

The hybrid model is estimated based on quarterly data on real Gross National Product (GNP), consumption, investment and hours worked, for the South African economy, over the period of 1970:1-2000:4. The model economy is a closed economy (i.e.  $Y_t = C_t + I_t$ ), where  $C_t$  and  $I_t$  are defined as final consumption expenditure by households and gross investment respectively<sup>3</sup>. The series are then converted into per-capita form by dividing them with the population aged by 15-64. Since there is no data for hours worked, we generate the series as follows. We assume employees work 40 hours per week and multiply it by the ratio of employment to the labor force.<sup>4</sup>

The hybrid model consisting of (5), (6), and (7) is in state-space form and can be estimated via a maximum likelihood approach. In our real business cycle model, output, consumption, and investment grow at the same rate of  $\eta$  in steady state. Before estimation, the series for output, consumption, and investment are de-trended by dividing with  $\eta$ . In addition, series for  $I_t$  is redundant in estimation since the resource constraint holds by construction in the data. Therefore,  $\hat{\pi}_t, \mu_t$ , and  $\xi_t$  is reduced to  $3 \times 1$  vector:

---

<sup>3</sup>Data are obtained from South African Reserve Bank Quarterly Bulletin, seasonally adjusted at constant price (2000 = 100).

<sup>4</sup>Data for employment is obtained from Statistics South Africa. Population aged 15-64 obtained from World Bank database is used as the proxy of labor force.



$$\hat{\pi}_t = [\hat{y}_t \ \hat{c}_t \ \hat{h}_t]'$$

$$\mu_t = [\mu_{yt} \ \mu_{ct} \ \mu_{ht}]'$$

$$\xi_t = [\xi_{yt} \ \xi_{ct} \ \xi_{ht}]'$$

and for all  $t = 1, 2, 3, \dots$ , the matrices  $D$  and  $V$  are:

$$D = \begin{bmatrix} d_{yy} & d_{yc} & d_{yh} \\ d_{cy} & d_{cc} & d_{ch} \\ d_{hy} & d_{hc} & d_{hh} \end{bmatrix}; \quad V = \begin{bmatrix} v_y^2 & v_{yc} & v_{yh} \\ v_{cy} & v_c^2 & v_{ch} \\ v_{hy} & v_{hc} & v_h^2 \end{bmatrix}$$

The structural parameters,  $\beta, \rho, \eta, \delta$ , and  $\psi$ , are constrained to satisfy the theoretical restrictions discussed in Section 2. The discount factor  $\beta$  and capital depreciation rate  $\delta$  are fixed in the estimation. The discount factor  $\beta$  is set equal to 0.99, as in Hansen (1985), which implies an annual real interest rate of four percent in steady state. The annual aggregate capital depreciation rate  $\delta$  is obtained from annual averaged values of  $\frac{I}{Y}$  and  $\frac{K}{Y}$ . This yields an annual depreciation rate of 0.076, or a quarterly rate of 0.019. The fixed  $\beta$  and  $\delta$  together with the estimated  $\rho, \eta, \gamma$ , and  $z$  help match the steady state values of  $y, c, h$  in the model with those in the data, whereas  $\psi$  and  $\sigma$  only affect the model's dynamics.

### 3.4 Results

In this section, we compare the out-of-sample forecasting performance of the hybrid DSGE-VAR model with the VARs, both Classical and Bayesian, in terms of the Root Mean Squared Errors (RMSEs). At this stage, a few words need to be said regarding the choice of the evaluation criterion for the out-of-sample forecasts generated from Bayesian models. As Zellner (1986: 494) points out “the optimal Bayesian forecasts will differ depending upon the loss function employed and the form of predictive probability density function”. In other words, Bayesian forecasts are sensitive to the choice of the measure used to evaluate the out-of-sample forecast errors. This fact was also observed in a recent study by Gupta (2006). However, Zellner (1986) points out that the use of the mean of the predictive probability density function for a series, is optimal relative to a squared error loss function and the Mean Squared Error (MSE), and, hence, the RMSE is an appropriate measure to evaluate performance of forecasts, when the mean of the predictive probability density function is used. This is exactly what we do below in Tables 8 through 11, when we use the average RMSEs over the one- to four-quarter-ahead forecasting horizon.

But, before we proceed to the discussion of the forecasting performance of the alternative models, it is important to lay out the basic structural differences and advantages of using BVARs over traditional VARs for forecasting.

### 3.4.1 Classical and Bayesian VARs

An unrestricted VAR model, as suggested by Sims (1980), can be written as follows:

$$\chi_t = C + \lambda(L)\chi_t + \varepsilon_t \quad (8)$$

where  $\chi$  is a  $(n \times 1)$  vector of variables being forecasted;  $\lambda(L)$  is a  $(n \times n)$  polynomial matrix in the backshift operator  $L$  with lag length  $p$ , i.e.,  $\lambda(L) = \lambda_1 L + \lambda_2 L^2 + \dots + \lambda_p L^p$ ;  $C$  is a  $(n \times 1)$  vector of constant terms; and  $\varepsilon$  is a  $(n \times 1)$  vector of white-noise error terms. The VAR model, thus, posits a set of relationships between the past lagged values of all variables and the current value of each variable in the model.

A crucial drawback of the VAR forecasts is “overfitting” due to the inclusion too many lags and too many variables, some of which may be insignificant. The problem of “overfitting” results in multicollinearity and loss of degrees of freedom, leads to inefficient estimates and large out-of-sample forecasting errors. Thus, it can be argued the performance of VAR forecasts will deteriorate rapidly as the forecasting horizon becomes longer.

A forecaster can overcome this “overfitting” problem by using Bayesian techniques. The motivation for the Bayesian analysis is based on the knowledge that more recent values of a variable are more likely to contain useful information about its future movements than older values. From a Bayesian perspective, the

exclusion restriction in the VAR is an inclusion of a coefficient without a prior probability distribution (Litterman, 1986a).

The Bayesian model proposed by Litterman (1981), Doan, *et al.* (1984), and Litterman (1986b), imposes restrictions on those coefficients by assuming they are more likely to be near zero. The restrictions are imposed by specifying normal prior<sup>5</sup> distributions with zero means and small standard deviations for all the coefficients with standard deviation decreasing as lag increases. One exception is that the mean of the first own lag of a variable is set equal to unity to reflect the assumption that own lags account for most of the variation of the given variable. To illustrate the Bayesian technique, suppose the “Minnesota prior” means and variances take the following form:

$$\begin{aligned}
 \beta_i &\sim N(1, \sigma_{\beta_i}^2) \\
 \beta_j &\sim N(0, \sigma_{\beta_j}^2)
 \end{aligned}
 \tag{9}$$

where  $\beta_i$  represents the coefficients associated with the lagged dependent variables in each equation of the VAR, while  $\beta_j$  represents coefficients other than  $\beta_i$ . The prior variances  $\sigma_{\beta_i}^2$  and  $\sigma_{\beta_j}^2$ , specify the uncertainty of the prior means,  $\beta_i = 1$  and  $\beta_j = 0$ , respectively.

---

<sup>5</sup>Note Litterman (1981) uses a diffuse prior for the constant, which is popularly referred to as the “Minnesota prior” due to its development at the University of Minnesota and the Federal Reserve bank at Minneapolis.

Doan *et al.* (1984) propose a formula to generate standard deviations as a function of a small number of *hyperparameters*<sup>6</sup> :  $w$ ,  $d$ , and a weighting matrix  $f(i, j)$ . This approach allows the forecaster to specify individual prior variances for a large number of coefficients based on only a few *hyperparameters*. The specification of the standard deviation of the distribution of the prior imposed on variable  $j$  in equation  $i$  at lag  $m$ , for all  $i, j$  and  $m$ , defined as  $S(i, j, m)$ :

$$S(i, j, m) = [w \times g(m) \times f(i, j)] \frac{\hat{\sigma}_i}{\hat{\sigma}_j} \quad (10)$$

where:

$$f(i, j) = \begin{cases} 1 & \text{if } i = j \\ k_{ij} & \text{otherwise, } 0 \leq k_{ij} \leq 1 \end{cases}$$

$$g(m) = m^{-d}, \quad d > 0$$

The term  $w$  is the measurement of standard deviation on the first own lag, which indicates the overall tightness. A decrease in the value of  $w$  results a tighter prior. The parameter  $g(m)$  measures the tightness on lag  $m$  relative to lag 1, and is assumed to have a harmonic shape with a decay of  $d$ . An increasing in  $d$ , tightens the prior as lag increases.<sup>7</sup> The parameter  $f(i, j)$  represents the tightness of variable  $j$  in equation  $i$  relative to variable  $i$ . Reducing the interaction parameter

---

<sup>6</sup>The name of *hyperparameter* is to distinguish it from the estimated coefficients, the parameters of the model itself.

<sup>7</sup>In this paper, we set the overall tightness parameter ( $w$ ) equal to 0.3, 0.2, and 0.1, and the harmonic lag decay parameter ( $d$ ) equal to 0.5, 1, and 2. These parameter values are chosen so that they are consistent with the ones that used by Liu and Gupta (2007).

$k_{ij}$  tightens the prior.  $\hat{\sigma}_i$  and  $\hat{\sigma}_j$  are the estimated standard errors of the univariate autoregression for variable  $i$  and  $j$  respectively. In the case of  $i \neq j$ , the standard deviations of the coefficients on lags are not scale invariant (Litterman, 1986b: 30). The ratio,  $\frac{\hat{\sigma}_i}{\hat{\sigma}_j}$  in (10), scales the variables so as to account for differences in the units of magnitudes of the variables.

The BVAR model is estimated using Theil’s (1971) mixed estimation technique, which involves supplementing the data with prior information on the distribution of the coefficients. For each restriction imposed on the parameter estimated, the number of observations and degrees of freedom are increased by one in an artificial way. Therefore, the loss of degrees of freedom associated with the unrestricted VAR is not a concern in the BVAR.

### 3.4.2 Forecast accuracy

Table 8 to 11 report the RMSEs from the hybrid DSGE-VAR model along with the VARs. The hybrid model does better job in predicting output and its components than it does in predicting hours worked.<sup>8</sup> To be more precise, for output and consumption the unconstrained hybrid model does better than the constrained hybrid model and the unrestricted VAR. However, for hours worked

---

<sup>8</sup>The hybrid model has 21 parameters, the six structural parameters  $\gamma, \rho, \eta, z, \psi$ , and  $\sigma$  from the real business cycle model, the fifteen elements from matrix D and V governing the behavior of the VAR residuals. For the constrained hybrid model, the number of parameters is reduced to 12. The VAR(1) model that we use to judge the hybrid model’s out-of-sample forecasting performance also has 21 parameters, output, consumption, and hours worked together with a constant and a linear time trend.

Table 8: **RMSE (2001Q1-2005Q4): Output**

QA	1	2	3	4	AVE
DSGE-VAR (Uncon)	<i>1.2432</i>	<i>1.7841</i>	<i>1.8214</i>	<i>1.7765</i>	<i>1.6563</i>
DSGE-VAR (Con)	1.3671	2.0595	2.4382	2.8072	2.1680
VAR (1)	1.4611	2.3092	2.8747	3.4087	2.5134
BVAR ( $w=.3, d=.5$ )	<b>0.6698</b>	<b>1.0454</b>	<b>1.3164</b>	<b>1.5712</b>	<b>1.1507</b>

QA: quarter ahead; RMSE: root mean squared error (%).

the constrained hybrid model outperforms the unconstrained one but not the unrestricted VAR. The scenario for investment is a bit different. The unconstrained hybrid model does better than the constrained one for only the one-quarter and two-quarters ahead out-of-sample forecasts, whereas for the three-quarters and four-quarters ahead forecasts the constrained hybrid model outperforms the unconstrained one.

As far as the forecasting performances of the BVARs are concerned, it is clear that the BVARs improve the out-of-sample forecast performance significantly. The RMSEs<sup>9</sup> generated from the BVARs are much smaller than those generated from both the hybrid model and the unrestricted VARs. In addition, the result suggests that a BVAR with a relatively loose prior produces smaller forecast errors. For all variables, output, consumption, investment and hours worked, a BVAR with the most loose prior ( $w = 0.3, d = 0.5$ ) performs the best.

---

<sup>9</sup>Here we only report the BVAR with the prior that does the best in terms of the out-of-sample forecasting performance.

Table 9: **RMSE (2001Q1-2005Q4): Consumption**

QA	1	2	3	4	AVE
DSGE-VAR (Uncon)	<i>1.2001</i>	<i>1.7884</i>	<i>2.1229</i>	<i>2.2967</i>	<i>1.8520</i>
DSGE-VAR (Con)	1.2287	1.9548	2.5158	3.0207	2.1800
VAR (1)	1.2029	1.7833	2.181	2.4643	1.9079
BVAR (w=.3, d=.5)	<b>0.5215</b>	<b>0.7080</b>	<b>0.8293</b>	<b>0.8570</b>	<b>0.7290</b>

QA: quarter ahead; RMSE: root mean squared error (%).

Table 10: **RMSE (2001Q1-2005Q4): Investment**

QA	1	2	3	4	AVE
DSGE-VAR (Uncon)	<i>2.8404</i>	<i>3.5985</i>	4.1179	4.1522	3.6773
DSGE-VAR (Con)	2.9518	3.6293	<i>3.9484</i>	<i>4.0228</i>	<i>3.6381</i>
VAR (1)	3.0437	4.3241	5.5072	6.4486	4.8309
BVAR (w=.3, d=.5)	<b>1.1230</b>	<b>1.4757</b>	<b>1.8097</b>	<b>2.0608</b>	<b>1.6173</b>

QA: quarter ahead; RMSE: root mean squared error (%).

Table 11: **RMSE (2001Q1-2005Q4): Hours worked**

QA	1	2	3	4	AVE
DSGE-VAR (Uncon)	2.5066	3.3475	4.0577	4.5857	3.6244
DSGE-VAR (Con)	<i>2.4477</i>	<i>2.9966</i>	<i>3.5075</i>	<i>3.7018</i>	<i>3.1634</i>
VAR (1)	2.3913	2.941	3.2884	3.2920	2.9782
BVAR5 (w=.3, d=.5)	<b>1.2420</b>	<b>1.6435</b>	<b>1.8927</b>	<b>1.9342</b>	<b>1.6781</b>

QA: quarter ahead; RMSE: root mean squared error (%).



In order to evaluate the models' forecast accuracy, we perform the across-model test between the hybrid model and the VAR(1), as well as the BVAR model. The across-model test is based on the statistic proposed by Diebold and Mariano (1995).<sup>10</sup> The across-model test results are reported in Table 12. The results indicate that, in general, the hybrid models outperform the unrestricted VAR(1) model for forecasting output and its components. One exception is consumption, the constrained hybrid model does not outperform the unrestricted VAR(1) model. However, most of these test statistics are not significant at 5% level. As far as the forecasting performance of the BVAR is concerned, the BVAR with the most loose prior ( $w = 0.3$ ,  $d = 0.5$ ) outperforms the hybrid models and the unrestricted VAR(1) model. In addition, most of these test statistics are significant either at 5% or 10% level. Finally, for hours worked, both constrained and unconstrained hybrid model do not outperform either the unrestricted VAR(1) model or the BVAR model, although few of the statistics are significant at 10% level.

---

<sup>10</sup>The test statistic is defined as the following. For instance, let  $\{e_t^v\}_{t=1}^T$  denote the associated forecast errors from the unrestricted VAR(1) model and  $\{e_t^h\}_{t=1}^T$  denote the forecast errors from the hybrid model. The test statistic is then defined as  $s = \frac{l}{\sigma_l}$ , where  $l$  is the sample mean of the "loss differentials",  $\{l_t\}_{t=1}^T$ , using  $l_t = (e_t^v)^2 - (e_t^h)^2$  for all  $t = 1, 2, 3, \dots, T$ , and where  $\sigma_l$  is the standard error of  $l$ . The  $s$  statistic is asymptotically distributed as a standard normal random variable and can be estimated under the null hypothesis of equal forecast accuracy, i.e.  $l = 0$ . Therefore, in this case, a positive value of  $s$  suggests that the hybrid model outperforms the unrestricted VAR(1) model in terms of out-of-sample forecasting.

Table 12: **Across-Model Test Statistics**

Quarters Ahead	1	2	3	4
(A) Output				
DSGE-VAR (Uncon) vs. VAR(1)	1.898*	1.823*	1.848*	1.616
DSGE-VAR (Con) vs. VAR(1)	0.888	1.718	1.592	1.576
DSGE-VAR (Uncon) vs. BVAR	-2.579**	-2.287**	-1.819*	-1.501
DSGE-VAR (Con) vs. BVAR	-2.740**	-2.310**	-1.916*	-1.598
VAR(1) vs. BVAR	-2.566**	-2.222**	-1.907*	-1.657
(B) Consumption				
DSGE-VAR (Uncon) vs. VAR(1)	0.126	-0.103	0.581	0.762
DSGE-VAR (Con) vs. VAR(1)	-0.418	-0.842	-1.055	-1.149
DSGE-VAR (Uncon) vs. BVAR	-3.267**	-1.935*	-1.643	-1.408
DSGE-VAR (Con) vs. BVAR	-3.760**	-2.035**	-1.859*	-1.499
VAR(1) vs. BVAR	-3.324**	-1.765*	-1.472	-1.197
(C) Investment				
DSGE-VAR (Uncon) vs. VAR(1)	0.604	0.985	1.093	1.166
DSGE-VAR (Con) vs. VAR(1)	0.329	1.568	1.633	1.466
DSGE-VAR (Uncon) vs. BVAR	-2.716**	-2.035**	-1.404	-1.148
DSGE-VAR (Con) vs. BVAR	-2.762**	-2.444**	-1.733*	-1.283
VAR(1) vs. BVAR	-2.394**	-2.086**	-1.605	-1.383
(D) Hours Worked				
DSGE-VAR (Uncon) vs. VAR(1)	-0.522	-0.915	-1.014	-0.976
DSGE-VAR (Con) vs. VAR(1)	-1.117	-0.717	-1.024	-1.208
DSGE-VAR (Uncon) vs. BVAR	-1.933*	-1.535	-1.321	-1.132
DSGE-VAR (Con) vs. BVAR	-1.947*	-1.686	-1.461	-1.345
VAR(1) vs. BVAR	-1.968*	-1.727*	-1.490	-1.329

Note: \* and \*\* indicate 10% and 5% significant respectively. BVAR is the optimal one with  $w = 0.3$  and  $d = 0.5$ .

### 3.5 Conclusion

In this paper, we develop an estimable DSGE model, in which we augment the linearized equations with a vector of residuals that follow a AR(1) process. The hybrid model, thus, combines the micro-founded DSGE model with the flexibility of the atheoretical VAR model, and hence, the name — DSGE-VAR. We then employ the hybrid model to measure the out-of-sample forecasting performance for output, consumption, investment, and hours worked for the South African economy over 2001:1-2005:4. The results indicate that, in general, the estimated hybrid DSGE model outperforms the Classical VAR, but not the Bayesian VARs. Moreover, the results suggest that a BVAR with a relatively loose prior produces smaller out-of-sample forecast errors.

The Hansen's (1985) version real business cycle model used in this paper is singular in the sense that the technology shock is the only shock to the system. Therefore, it is necessary to study the importance of various shocks in accounting for the dynamic behaviour of output and its main components. In this regard, future research aims to estimate a New Keynesian DSGE model, which will allow us to incorporate nominal shocks. Further, we also aim to estimate the current model using Bayesian techniques. The ultimate goal of all these future extensions will be to analyze whether the DSGE model can outperform the BVARs, as far as forecasting is concerned.

## A. Optimization

In our model economy, the representative consumer problem is to maximize the utility function (1) by choosing  $\{C_t, H_t, K_{t+1}\}_{t=0}^{\infty}$

$$U(C_t, H_t) = E_t \sum_{t=0}^{\infty} \beta^t (\ln C_t - \gamma H_t), \quad 0 < \beta < 1 \quad \gamma > 0$$

subject to the resource constraint:

$$Z_t K_t^\rho (\eta^t H_t)^{1-\rho} \geq C_t + K_{t+1} - (1 - \delta)K_t \tag{A.1}$$

From the (A.1), we have:

$$C_t = Z_t K_t^\rho (\eta^t H_t)^{1-\rho} + (1 - \delta)K_t - K_{t+1} \tag{A.2}$$

The Bellman equation for this problem:

$$V(K_t, Z_t) = \max_{H_t, K_{t+1}} \{ \ln [Z_t K_t^\rho (\eta^t H_t)^{1-\rho} + (1 - \delta)K_t - K_{t+1}] - \gamma H_t \} + \beta E_t V(K_{t+1}, Z_{t+1}) \tag{A.3}$$

The first order condition (FOC) for hours worked:

$$\frac{\partial V(K_t, Z_t)}{\partial H_t} = 0 \tag{A.4}$$

$$\frac{1}{C_t} (1 - \rho) Z_t K_t^\rho \eta^{t(1-\rho)} H_t^{-\rho} - \gamma = 0 \tag{A.5}$$

$$\gamma = \frac{Y_t}{C_t} (1 - \rho) \frac{1}{H_t} \tag{A.6}$$

and the FOC for capital stock:

$$\frac{\partial V(K_t, Z_t)}{\partial K_{t+1}} = 0 \quad (\text{A.7})$$

$$\frac{1}{C_t}(-1) + \beta E_t V(K_{t+1}, Z_{t+1}) = 0 \quad (\text{A.8})$$

The envelope condition is:

$$\frac{\partial V(K_t, Z_t)}{\partial K_t} = \frac{1}{C_t} Z_t \rho K_t^{\rho-1} (\eta^t H_t)^{1-\rho} + (1 - \delta) \quad (\text{A.9})$$

Updating (A.9) and combining with (A.8) yields the Euler equation for capital stock:

$$\frac{1}{C_t} = \beta E_t \left[ \frac{1}{C_{t+1}} Z_{t+1} \rho K_{t+1}^{\rho-1} (\eta^{t+1} H_{t+1})^{1-\rho} + (1 - \delta) \right] \quad (\text{A.10})$$

$$\frac{1}{C_t} = \beta E_t \left\{ \frac{1}{C_{t+1}} \left[ \frac{Y_{t+1}}{K_{t+1}} \rho + (1 - \delta) \right] \right\} \quad (\text{A.11})$$

## B. The steady state and log-linearization

### B.1 The steady state

The complete model economy:

$$Y_t = C_t + I_t \quad (\text{B.1})$$

$$Y_t = Z_t K_t^\rho (\eta^t H_t)^{1-\rho} \quad (\text{B.2})$$

$$K_{t+1} = (1 - \delta)K_t + I_t \quad (\text{B.3})$$

$$\gamma = \frac{Y_t}{C_t} (1 - \rho) \frac{1}{H_t} \quad (\text{B.4})$$

$$\frac{1}{C_t} = \beta E_t \left\{ \frac{1}{C_{t+1}} \left[ \left( \frac{Y_{t+1}}{K_{t+1}} \right)^\rho + (1 - \delta) \right] \right\} \quad (\text{B.5})$$

$$\log Z_t = (1 - \psi) \log \bar{Z} + \psi \log Z_{t-1} + \epsilon_t, \quad \epsilon_t \sim i.i.d.(0, \sigma_\epsilon^2) \quad (\text{B.6})$$

In equilibrium,  $y_t = Y_t/\eta^t$ ,  $c_t = C_t/\eta^t$ ,  $i_t = I_t/\eta^t$ ,  $h_t = H_t$ ,  $k_t = K_t/\eta^t$ , and  $z_t = Z_t$ , therefore we can rewrite the model as:

$$y_t = c_t + i_t$$

$$y_t = z_t k_t^\rho h_t^{1-\rho}$$

$$\eta k_{t+1} = (1 - \delta)k_t + i_t$$

$$\gamma = \frac{y_t}{c_t} (1 - \rho) \frac{1}{h_t}$$

$$\frac{\eta}{c_t} = \beta E_t \left\{ \frac{1}{c_{t+1}} \left[ \left( \frac{y_{t+1}}{k_{t+1}} \right)^\rho + (1 - \delta) \right] \right\}$$

$$\log z_t = (1 - \psi) \log \bar{Z} + \psi \log z_{t-1} + \epsilon_t, \quad \epsilon_t \sim i.i.d.(0, \sigma_\epsilon^2)$$

In steady state we have  $y_t = y$ ,  $c_t = c$ ,  $i_t = i$ ,  $h_t = h$ ,  $k_t = k$ , and  $z_t = z$  for all  $t = 0, 1, 2, \dots$ . Solving for the steady state values of the six variables:

$$a = A \tag{B.7}$$

$$k = \left( \frac{\rho}{\eta/\beta - 1 + \delta} \right) y \tag{B.8}$$

$$i = \left[ \frac{\rho(\eta - 1 + \delta)}{\eta/\beta - 1 + \delta} \right] y \tag{B.9}$$

$$c = \left\{ 1 - \left[ \frac{\rho(\eta - 1 + \delta)}{\eta/\beta - 1 + \delta} \right] \right\} y \tag{B.10}$$

$$h = \left( \frac{1 - \rho}{\gamma} \right) \left\{ 1 - \left[ \frac{\rho(\eta - 1 + \delta)}{\eta/\beta - 1 + \delta} \right] \right\}^{-1} \tag{B.11}$$

$$y = z^{1/(1-\rho)} \left( \frac{\rho}{\eta/\beta - 1 + \delta} \right)^{\rho/(1-\rho)} \left( \frac{1 - \rho}{\gamma} \right) \left\{ 1 - \left[ \frac{\rho(\eta - 1 + \delta)}{\eta/\beta - 1 + \delta} \right] \right\}^{-1} \tag{B.12}$$

## B.2 Log-linearization

This section presents the log-linearized DSGE model. The principle of log-linearization is to replace all equations by Taylor approximation around the steady state, which are linear functions in the log-deviations of the variables (Uhlig, 1995:4). Suppose  $\Pi_t$  be the vector of variables,  $\pi$  their steady state, and  $\hat{\pi}_t$  the vector of log-deviations:

$$\hat{\pi}_t = \log \Pi_t - \log \pi \tag{B.13}$$

in other words,  $\hat{\pi}_t$  denote the percentage deviations from their steady state levels.

Using first-order Taylor approximations to rewrite all the equations of the model:

$$\hat{y}_t = \hat{z}_t + \rho \hat{k}_t + (1 - \rho) \hat{h}_t \quad (\text{B.14})$$

$$\hat{z}_t = \psi \hat{z}_{t-1} + \varepsilon_t \quad (\text{B.15})$$

$$\left(\frac{\eta}{\beta} - 1 + \delta\right) \hat{y}_t = \left[\left(\frac{\eta}{\beta} - 1 + \delta\right) - \rho(\eta - 1 + \delta)\right] \hat{c}_t + \rho(\eta - 1 + \delta) \hat{i}_t \quad (\text{B.16})$$

$$\eta \hat{k}_{t+1} = (1 - \delta) \hat{k}_t + (\eta - 1 + \delta) \hat{i}_t \quad (\text{B.17})$$

$$\hat{c}_t + \hat{h}_t = \hat{y}_t \quad (\text{B.18})$$

$$\begin{aligned} 0 = & \frac{\eta}{\beta} \hat{c}_t - \frac{\eta}{\beta} E_t \hat{c}_{t+1} + \left(\frac{\eta}{\beta} - 1 + \delta\right) E_t \hat{y}_{t+1} \\ & - \left(\frac{\eta}{\beta} - 1 + \delta\right) \hat{k}_{t+1} \end{aligned} \quad (\text{B.19})$$